fies the level of trust one can have with respect to its announcing valid updates. To compute the reputation, AS-CRED analyzes the past updates announced by each observable AS in the Internet, over a time-window, based on well-defined properties. It then classifies the resulting observations into multiple types of feedback. The feedback values are input into a mathematical function for computing AS reputation. The reputation is then used to track the instances of invalid updates announced in the Internet and trigger *alerts*. The **contributions** of the paper are: (1) a reputation service for ASes, characterizing their trustworthiness; (2) a set of well-defined properties for analyzing AS behavior; (3) a simple reputation function and feedback mechanism; (4) a reputation portal which regularly publishes AS reputation; and (5) a reputation-based alert service which tracks potentially invalid updates in the Internet. Detailed analysis of AS-CRED demonstrates: (a) AS behavior is repetitive making reputation an effective trust metric, and (b) AS-CRED's alerts for invalid updates show an eight fold improvement over existing alert systems.

AS-CRED: Reputation Service for Trustworthy Inter-domain Routing

Jian Chang* Krishna K. Venkatasubramanian† Andrew G. West‡

Sampath Kannan** Insup Lee††

Boon Thau Loo‡‡ Oleg Sokolsky§

*University of Pennsylvania
†University of Pennsylvania
‡University of Pennsylvania
**University of Pennsylvania, kannan@cis.upenn.edu
††University of Pennsylvania, lee@cis.upenn.edu
‡‡University of Pennsylvania, boonloo@cis.upenn.edu
§University of Pennsylvania, sokolsky@cis.upenn.edu

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-10-17.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE
AS-CRED: Reputation Service for Trustworthy Inter-domain Routing

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pennsylvania,Department of Computer and Information Science,Philadelphia,PA,19104

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**12**	

AS-CRED: Reputation Service for Trustworthy Inter-domain Routing*

Jian Chang, Krishna K. Venkatasubramanian, Andrew G. West, Sampath Kannan,
Insup Lee, Boon Thau Loo and Oleg Sokolsky
Department of Computer and Information Science,
University of Pennsylvania, Philadelphia, PA, 19104
{jianchan, vkris, westand, kannan, lee, boonloo, sokolsky}@cis.upenn.edu

ABSTRACT

The current design of BGP implicitly assumes the existence of trust between ASes with respect to exchanging valid BGP updates. This assumption of complete trust is problematic given the frequent announcement of invalid — inaccurate or unnecessary — updates. This paper presents *AS-CRED*, a reputation service for ASes which quantifies the level of trust one can have with respect to its announcing valid updates. To compute the reputation, AS-CRED analyzes the past updates announced by each observable AS in the Internet, over a time-window, based on well-defined properties. It then classifies the resulting observations into multiple types of feedback. The feedback values are input into a mathematical function for computing AS reputation. The reputation is then used to track the instances of invalid updates announced in the Internet and trigger *alerts*. The **contributions** of the paper are: (1) a reputation service for ASes, characterizing their trustworthiness; (2) a set of well-defined properties for analyzing AS behavior; (3) a simple reputation function and feedback mechanism; (4) a reputation portal which regularly publishes AS reputation; and (5) a reputation-based alert service which tracks potentially invalid updates in the Internet. Detailed analysis of AS-CRED demonstrates: (a) AS behavior is repetitive making reputation an effective trust metric, and (b) AS-CRED's alerts for invalid updates show an eight fold improvement over existing alert systems.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols; C.2.3 [**Computer-Communication Networks**]: Network Operations

General Terms

BGP, autonomous systems, trust management, reputation, alert service

1. INTRODUCTION

The Border Gateway Protocol (BGP) is the standard communication protocol for interconnecting large IP domains, called Autonomous Systems (AS). BGP operates by exchanging *updates* between ASes; which contains reachability information for prefixes (IP address blocks). The current design of BGP implicitly requires the existence of complete trust between ASes exchanging BGP routing information. In reality, this implicit assumption of complete trust is a cause for concern as many ASes announce *invalid* updates for some or all of their prefixes.

The invalidity of updates in the literature has usually referred to prefix hijacking — announcing the reachability to a prefix which the AS does not own[1] [17, 19]. Over the past decade numerous cases of such inaccurate updates have been documented. For example, AS7007 incident in 1997 [1] and invalid announcement of prefixes belonging to eBay by AS10139 [22]. However, our experiments with actual BGP data revealed the existence of an additional type of invalid updates, where the prefix being announced is repeatedly unstable. That is, ASes announce and withdraw prefixes they own in quick succession, effectively making them useless for data traffic. For example, AS37035 was seen announcing and withdrawing the prefix 41.222.179.0/24, which it owns, 4824 times between Dec. 3, 2009 and Dec. 7, 2009. This amounts to announcing (and withdrawing) the prefix repeatedly once every 1.5 minutes, on average. Such unnecessary updates increase the burden on the Internet where the routers are already heavily loaded given its rapid growth [20].

We argue that it is useful to *quantify the level of trust* one can have on an AS with respect to announcing valid updates. We define a *valid* update as satisfying two conditions: (1) it provides *accurate* routing information, *i.e.*, no prefix hijacking; and (2) the update itself is *necessary* for the correct operation of the Internet, *i.e.*, it is not part of a sequence of short duration prefix announcements and withdrawals. Much work has been done in detecting occurrence of invalid updates in the Internet [13, 18, 17, 19, 21, 24]. These solutions however are limited to detecting inaccurate updates, none of them are designed to address the *necessity* aspect of update validity. In this paper, we present *AS-CRED*, a reputation management service for Autonomous Systems. It quantifies the level of "cred" (trust) one can

*This research was supported in part by ONR MURI N00014-07-1-0907. POC: Insup Lee, lee@cis.upenn.edu

[1] We use the term *own* to describe prefixes allocated to ASes by Internet address registries such as IANA, or those belonging to their customers.

have in an AS' tendency to announce *valid updates.* Trust in AS-CRED is represented using a predictive metric called *reputation.* To compute the reputation of an AS, AS-CRED analyzes its past updates received, over a time-window based on well-defined properties. Out of this analysis: it creates a *white-list* of AS-prefix pairs which it considers legitimate, and it provides feedback to a reputation function to compute the *reputation value* of all the observable ASes in the Internet. The reputation value and white-list are then used to track the instances of inaccurate and unnecessary updates announced over the Internet and trigger *alerts.* A comparison of the alerts with well-known prefix hijacking alert system Internet Alert Registry (IAR) [6] showed that AS-CRED reduces the number of false positives (valid updates flagged as invalid) by about eight fold. The analysis time-window is shifted and reputation recomputed on a daily basis, making it a dynamic value. Moreover, as BGP function by exchanging reachability information about all the active ASes and prefixes in the Internet, AS-CRED can be used to compute reputation values for the entire Internet at the inter-domain level.

AS-CRED service has many uses: (1) **Behavior Metric:** Its association of an objective and global trust metric with every observable AS in the Internet allows ASes to not only know about other ASes but also how it itself is perceived. AS can now make better informed decisions in dealing with others and tuning their business, traffic, scalability or security policies, accordingly; (2) **White-List:** One of the byproducts of reputation computation is a white-list of AS-prefix pairs which are legitimate (stable and legal). The white-list can be used by ASes for tuning their import and export policies; (3) **Expanded Alert Service:** The alert mechanism is unlike any existing alert systems available [6, 19], in that: (a) it provides an alert for both inaccurate and unnecessary updates announced, (b) it provides the reputation value for the AS involved along with the alert, which is very useful for understanding the behavior of ASes, and (c) the reputation and alerts can provide effective diagnostic and forensic tool to debug network connectivity issues at Internet scale; and (4) **Incentivization:** The availability of reputation has the potential to provide an incentive for ASes to improve their behavior in the future.

The **contributions** of the paper are: (1) a reputation service for ASes, characterizing their trustworthiness; (2) a set of well-defined properties for analyzing AS behavior; (3) a simple reputation function and feedback mechanism; (4) a reputation portal which regularly publishes AS reputation; and (5) a reputation-based alert service which tracks potentially invalid updates in the Internet. Detailed analysis of AS-CRED demonstrates: (a) AS behavior is repetitive making reputation an effective trust metric, and (b) AS-CRED's alerts for invalid updates show an eight fold improvement over existing alert system [6].

The paper is organized as follows; Section 2 presents the background. Section 3 presents details of AS-CRED architecture and data source. Section 4 presents the feedback and reputation model. Section 5 presents the AS-behavior analysis and reputation usage of AS-CRED. Section 6 presents the results of our analysis of AS-CRED and performance results. Section 7 presents the related work. Finally, Section 8 concludes the paper.

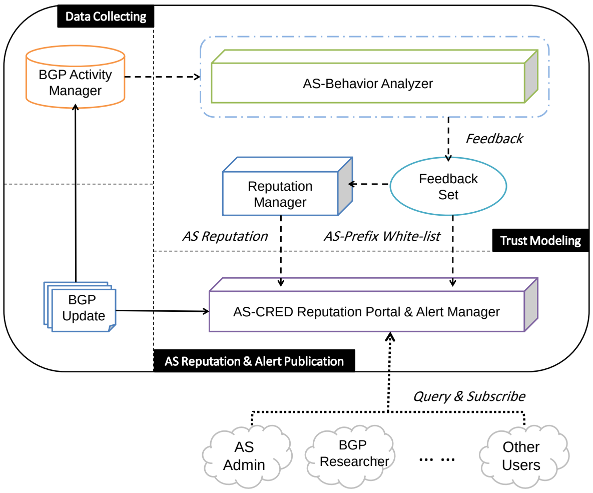


Figure 1: AS-CRED Architecture

2. BACKGROUND

In this section, we present a brief overview of the Border Gateway Protocol and the problem statement.

2.1 The Border Gateway Protocol

Border Gateway Protocol is a path-vector routing protocol for exchanging information about reaching IP address blocks (prefixes). Using BGP, an AS X, which owns a prefix p, originates an *update* notifying its neighboring AS Y of its ownership. AS Y then forwards this update further to its neighbor AS Z by adding its own AS number to the path vector, called AS_PATH, in the update. This informs AS Z that in order to reach the prefix p, the gateway router at AS Y is the next hop. When an update is received at an AS (at its BGP router), it determines whether the update should be accepted or not. The acceptance of an update means that the router is willing to add the route to the prefix, to its routing table. Note that BGP itself does not have any mechanism for making decisions regarding route preference. Its only purpose is to convey the reachability of prefixes to ASes. Each AS has its own policies that determine whether it accepts a BGP update and whether it forwards it (update) to its neighbors. Routing policies serve an important purpose in BGP and provide an AS with not only the capability to prefer routes over others to reach a prefix, but also to filter and/or tag update to change the route's relative preference downstream. In terms of implementation, AS policies are of three types - *import, decision process, and export.* Import policies are used to determine which routes to accept. Decision policies are then applied to the imported routes to choose the best one for each prefix. Finally, export policies are used to determine which neighbors get to know about a particular route [11].

2.2 Problem Statement

The current version of BGP [2] was designed with only effectiveness in mind. It implicitly assumes ASes can be trusted to announce valid updates. An update is considered *valid,* if it satisfies two conditions: (1) *Accuracy:* the information in the update regarding announcement (or withdrawal) of

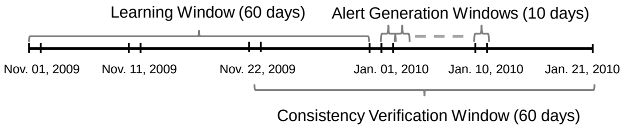

Learning Window (60 days)
Alert Generation Windows (10 days)
Nov. 01, 2009
Nov. 11, 2009
Nov. 22, 2009
Jan. 01, 2010
Jan. 10, 2010
Jan. 21, 2010
Consistency Verification Window (60 days)

Figure 2: Data Source Time Windows

a prefix is accurate. That is, the AS does not include incorrect announcer or prefix information; and (2) *Necessity:* the update is necessary for ASes in the Internet to be able to reach the prefix contained within, in a sustained manner. An update is considered *invalid* if it does not satisfy *at least one* of these two conditions. *The principal* question being addressed in this paper is, how to develop a trust metric for ASes which characterizes their tendencies to announce valid updates. In this work we assume that malicious ASes do not arbitrarily modify the ID of the announcer of an update. We believe that this assumption does not necessarily reduce the importance of the problem being addressed as such attacks can be only be mitigated by cryptographic approaches such as S-BGP [18]. Moreover, this is a problem faced by all schemes which deal with BGP update semantics, such as prefix hijack detection, and is not unique to us.

3. REPUTATION FOR AUTONOMOUS SYSTEMS

In this paper, we present *AS-CRED*, a service for computing and associating a *trust* value with ASes in the Internet. In general, trust is defined as the subjective expectation in the competence of an entity to act dependably within a particular context [14]. Here, the context in question is announcing valid updates. One of the well known ways to quantify trust is to use the notion of *reputation.* Reputation of an entity is a characterization of its past in performing a specific task. For entities that preserve their behavior, reputation forms an effective predictive model. In order to use reputation successfully, there are three prerequisites: (1) identifying behaviors of interest; (2) monitoring for exhibition of the behaviors; and (3) providing feedback on the experience. Once the feedback has been received, the reputation can be computed based on a well-defined mathematical function.

3.1 AS-CRED Architecture

Figure 1 shows the AS-CRED architecture. Its principal task is to maintain reputation for every observable AS in the Internet. AS administrators can query this database of reputation to determine the extent of trust they can place on the ASes of interest to them. AS-CRED has four main components: (1) **BGP Activity Manager:** This is a database which collects latest BGP updates from well-connected BGP data collectors (*e.g.*, RouteViews). The data provides a view of all active ASes in the Internet and the prefixes that they announce at different times; (2) **Reputation Manager:** This computes the reputation of the ASes based on a well-defined mathematical function using past behavior information in the form of feedback; (3) **AS-Behavior Analyzer:** This component analyzes the updates managed by the BGP Activity Manager, within a specific time duration called the *learning window*, based on a set of well-defined properties. The results of the analysis, which is a classification of the past behavior of ASes, forms the feedback for the Reputation Manager; (4) **Reputation Portal:** Once the AS reputations are computed we have a (cumulative) value of their trustworthiness, which is made available through a web portal; and (5) **Alert Manager:** AS-CRED has the knowledge of prefixes announced by different ASes over time, along with their validity. It therefore uses this information, along with AS reputation, to trigger real-time alerts regarding potential invalidity of any new updates propagated within the Internet.

Note that, AS-CRED dynamically updates reputation of ASes as their behavior changes over time. In this regard, the AS-Behavior Analyzer continuously evaluates the updates received over a sliding window (which includes newer updates and excludes old ones at the beginning of the window), and provides updated feedback for the Reputation Manager. In the next three sections, we present these principal components of AS-CRED architecture in detail. We now present a brief description of the data source used for populating the BGP Activity Manager.

3.2 Data Source

AS-CRED depends upon BGP updates to compute AS reputation. We use the RouteViews BGP trace collector maintained by University of Oregon [9], to populate the BGP Activity Manager. The RouteViews trace collector is a group of BGP routers which peer with a large number of ISPs via BGP sessions. At the time of writing, the RouteViews BGP trace collector received BGP updates from 46 ASes. It has been shown in [25] that RouteViews covers almost all the ASes currently active within the Internet and is therefore a good source for computing reputation of ASes. In this work, we use data from Nov. 1, 2009 - Dec. 31, 2009 (see Figure 2). This 60 day[2] period is the *learning window*, based on which AS behavior is analyzed and reputation is computed. We wanted the learning window to be sufficiently long that it is not biased by any transient AS behavior, more justification for the choice is presented in Section 6.2. The reputation of ASes is then used to generate alerts for the invalid updates received on Jan. 1, 2010. The learning window is then moved forward by one day now covering the days from Nov. 2, 2009 to Jan. 1, 2010, and alerts are generated for updates received on Jan. 2, 2010. In this manner, we compute reputation and generate alerts for 10 days covering Jan. 1, 2010 to Jan. 10, 2010. We call each day in this 10 day period *alert generation window.* We find that recomputing reputations *once a day* is sufficient from both a complexity and predictive capability standpoint. The choice of 10 days was to demonstrate the AS reputation trend over time. We also define a third window called the *consistency verification window* for purely analysis purposes. The consistency verification window is used for verifying if the computed reputation is a good representation of the behavior analysis. The consistency verification window is 60 days long and includes both the 10 alert generation windows and 10 additional days. The idea is to allow sufficient time for the AS-prefix pairs, received after the learning window, to evolve and therefore be analyzable with the benefit of hindsight.

4. AS REPUTATION MANAGEMENT

[2] In order to be fair, we did not consider updates announced within 24 hours of the end of the learning window in computing the reputation of the ASes as such bindings have not had enough time to prove themselves.

Table 1: AS-prefix Binding Stability for Documented Prefix Hijacking Instances

Date	Prefix Hijacked	Victim AS	Attacker AS	Duration	# of Occurrence
Jan. 13, 2007	12.0.0.0/8	7018	31604	4 hours 26 minutes	1
Feb. 24, 2008	208.65.153.0/24	36561 (YouTube)	17557	9 hours 45 minutes	1
March 15, 2008	194.9.82.0/24	36915	6461	17 minutes	1
Dec. 2004 - Jan. 2005	61.0.0.0/8	NULL	4787	< 1 minute	100+
Dec. 2004 - Jan. 2005	82.0.0.0/8	NULL	8717	< 1 minute	100+

The reputation value assigned by AS-CRED to ASes is actually designed to characterize their "untrustworthiness" in announcing valid updates. The reason for using reputation to capture untrustworthiness is that ASes announcing valid updates far outnumber those that do not. That the Internet functions flawlessly most of the time is an attestation to this fact. However, invalid updates have considerable negative influence on the operations of large portions of the Internet. A case in point is the invalid update announced by an AS belonging to Pakistan Telecom, which altered the routes to IP prefixes belonging to YouTube. This announcement reportedly resulted in more than two-thirds of the Internet not being able to access YouTube [7]. In this section, we describe the types of feedback and the reputation function used by the Reputation Manager.

4.1 Feedback Types

A feedback for AS-CRED is a tuples of the form $\{a, p, t\}$, where a is the AS announcing the prefix p at time-stamp t. It can be of one of three types: (1) **Good:** This feedback is provided each time an AS announces a required update with accurate routing information. The AS and prefix involved are referred to as exhibiting *good behavior* and are stored in a set G where every tuple is of the form $g_i = \{a, p, t\}$; (2) **Unnecessary (Bad):** This feedback is provided each time an AS accidentally announces *unnecessary* update. The AS and prefix involved are stored in a set B where every tuple is of the form $b_i = \{a, p, t\}$; and (3) **Inaccurate (Ugly):** This feedback is provided each time an AS announces an *inaccurate* routing information. The AS and prefix involved are stored in a set U where every tuple is of the form $u_i = \{a, p, t\}$.

The *Good* set keeps track of the valid prefixes announced by ASes within the learning window. The *Bad* and the *Ugly* set keep track of prefixes announced by ASes which are invalid as a result of being unnecessary and inaccurate, respectively. In the rest of the paper, we use the term *GBU sets* to refer to the three feedback sets, collectively. The act of announcing unnecessary or inaccurate updates are collectively called *poor behaviors*. The GBU sets form the feedback that is provided for AS reputation computation. Note that, an AS may demonstrate good behavior for one prefix but simultaneously demonstrate poor behaviors for others. Also, at any given time, a particular AS-prefix pair is exclusively classified in one of the three feedback types. Finally, as the feedbacks are generated locally, we do not have to consider the case of potentially dishonest feedback affecting our reputation computation outcome.

4.2 AS Reputation Function

The reputation function computes the reputation of an AS based on the feedback in the GBU sets. As reputation is being used to characterize the untrustworthiness of an AS in announcing valid updates, the reputation computed for it has three properties: (1) the initial reputation of the AS is always maximum and decreases as incidences of poor behaviors are observed; (2) more recently observed poor behaviors are weighed more heavily than older ones, as it has been observed that a recent poor behavior is usually a precursor to another one; and (3) the reputation of an AS a is a vector of the form $[Rep_B(a), Rep_U(a)]$, where $Rep_B(a)$ is the reputation of an AS a based on each of its entry in the B set. Similarly, $Rep_U(a)$ is the reputation of an AS a based on each of its entry in the U set.

Table 2: Prevalence Persistence and Feedback

Prevalence	Persistence	Feedback
high	high	Good (G)
high	low	Unnecessary (B)
low	high	Good (G)
low	low	Inaccurate (U)

Formally, reputation is defined by the following function:

$$Rep_X(a) = \sum_{t_i} 2^{-(t_{now}-t_i)/h_X} \qquad (1)$$

Here $Rep_X(a)$ is the reputation of an AS a for exhibiting one of the poor behavior X, t_{now} is the current time and $t_i \in X.T$ is the time-stamp of when a X was observed, and h_x is the half-life of the decay function for exhibiting the behavior X. The values of $t_{now} - t_i$ are in the same units as h. It can be seen that the reputation returned for an AS varies between 0 (excellent) and REP_MIN[3] (poor). To set the half-life values for $Rep_U(a)$ and $Rep_B(a)$, we calculate the average time difference between two entries in the U and B set, for each AS. Based on this, we set the half-life h_U and h_B to values within which a large majority of the ASes repeat the respective behavior at least once. As all ASes are considered "innocent until proven guilty", they will all have a initial reputation vector of [0,0], which will change depending upon their appearance in the B or U set. The most important aspect of computing reputation of an AS is to be able to observe and provide feedback on the validity of updates announced by it. We now describe the properties that we use for evaluating the validity of updates.

5. AS-BEHAVIOR ANALYSIS AND ALERT GENERATION

In this section, we present a detailed look at the set of properties which are used by the AS Behavior Analyzer component of AS-CRED for providing feedback for reputation computation. In this regard, we consider the following two

[3] The absolute worst AS is the one which has an entry in the B or U sets for every possible time-stamp in the learning window and at each time-stamp it has committed a poor behavior for all possible prefixes in the IP address space. Therefore, $REP\_MIN = \int 2^{-\frac{(t_{now}-t_i)}{h}} \times \sum_{i=1}^{32} 2^i$.

Table 3: AS-CRED (AS-Behavior Analysis) vs. IAR w.r.t. IRR (NR: No Record in IRR, MR: Match in IRR (FP), NMR: No Match in IRR (Hijack))

Scheme	NR	MR	NMR
AS-CRED	841 (13.7%)	975 (18.4%)	4323 (81.6%)
IAR	4190 (10.7%)	25892 (74.4%)	8903 (25.6%)

properties: *AS-prefix value legality* and *AS-prefix binding stability*. Before we proceed, we define the term *AS-prefix binding* as a claimed ownership of a particular prefix by an AS. It is established when an AS announces a prefix for the first time. A binding may have many *incidences*, which is defined as an announcement and corresponding withdrawal of a prefix by an AS. An AS-prefix binding *terminates*, when an AS withdraws the prefix and never announces it again. In the rest of the paper we use the terms AS-prefix binding, prefix binding and binding, interchangeably. We begin by describing the various behavior analysis properties used by AS-CRED to provide feedback along with its capabilities. We then move on to describing the how the reputation information is disseminated and used.

5.1 Property I: AS-prefix Value Legality

One of the most important properties of AS-prefix binding, for determining the validity of an update, is the legality of the values. Typically, ASes and prefixes can only take a range of values. Any value which does not fall under this range is considered illegal. Both the values of ASes and prefixes can be illegal. For example, an AS number is illegal if its value is in the range of 64496-64511 (reserved for use in documentation and sample code), 64512-65534 (designated for private use) or 65535 (reserved) [3]. Similarly, an AS could claim to own a prefix that has not been assigned to anyone. Such prefixes, called *bogons* are spread over the entire IP address range and can be found in bogon lists maintained by IANA [5].

One of the most common reasons for illegal AS numbers to appear in updates is the incorrect application of export policies by BGP gateway routers [20]. This allows an illegal AS number to remain in the update as it is forwarded to neighboring ASes. We classify such cases into the B set due to the *nonnecessity* of private AS numbers in the updates. However, as the AS number in question is not legal, which AS do we penalize? Given the fact that we observe illegal AS numbers in the Internet due to lack of proper filtering, we penalize the first AS which has a legitimate AS number in the AS path. Therefore, given an update of the form $\{p, AS\_PATH =< AS_a, AS_b, ..., AS_n, AS_x >\}$, where AS_x is the announcer of the prefix p, and AS_a is the neighbor of RouteViews. If AS_x has a illegal value and AS_n onward all others have legal ones, we add an entry of the form $\{AS_n, p, t\}$ to the B set. Here, t is the time-stamp of when the update was received. On the other hand, by announcing bogons, ASes are providing inaccurate information about their prefix ownership. Further, an AS might want to announce an heretofore unallocated prefix for spamming purposes as suggested in [23]. Therefore, ASes which send bogons are added to the U set, using an entry of the form $\{AS_m, -, t\}$. Here, AS_m is the AS announcing the bogon, t is time-stamp of when the update was received. We leave the entry for the prefix blank as it is not a valid one.

Table 4: Examples of Unnecessary BGP Updates (NAW: Number of Announcements and Withdrawals)

AS	Prefix	NAW	Duration Observed
8452	41.235.83.0./24	2088	Nov. 2 - Nov. 10, 2009
704	152.63.49.180/30	1628	Dec. 8 - Dec. 31, 2009
145	140.217.157.0/24	1080	Nov. 1 - Nov. 27, 2009

The legality properties are amenable to checking based on largely static lists. This allows us to use them for providing real-time feedback to the Reputation Manager.

5.2 Property II: AS-prefix Binding Stability

In the inter-domain routing world, we find that legitimate AS-prefix bindings last for long durations of time and are very stable in nature [10, 17]. On the other hand, lower binding duration implies greater chances of invalid updates. Table 1 shows a list of well-known hijacked prefix announcements in the past. We find that all of them last for a very short duration of time, *i.e.*, have very low stability. Inspired by this observation, we present two metrics which can be used to compute the level of stability of AS-prefix bindings and can therefore be used to potentially deduce nonnecessity or inaccuracy of the updates which contain them.

5.2.1 Prevalence and Persistence

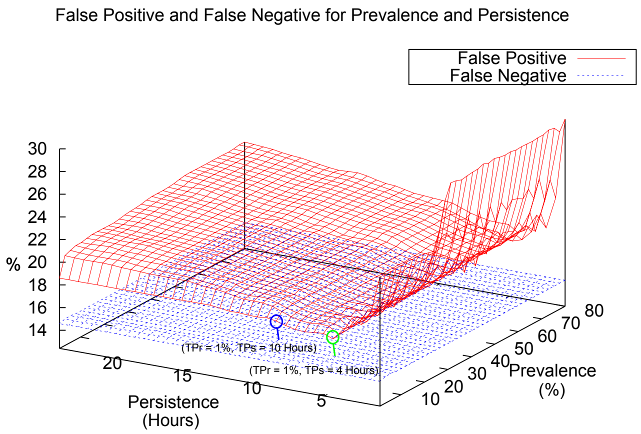
Prevalence (Pr) of a AS-prefix binding is the percentage of time a prefix is claimed to be reachable by an AS within a time window (the learning window in our case). More formally:

$$Pr(p, AS) = \sum_i (Tw^i(p, M) - To^i(p, M))/T_{learn} \qquad (2)$$

Here, i is the index of all the announcements of prefix p by AS M during T_{learn} (the learning window). $Tw(p, M)$ is the time prefix p is the withdrawn by M. $To(p, M)$ is the time prefix p is the announced by M. If the prevalence is above a threshold then the binding is considered stable. However, the prevalence metric alone is not sufficient, as it will not be able to detect instability due to repeated short-duration binding incidences, that is, AS-prefix bindings which are unnecessary. We therefore consider another metric in conjunction with prevalence, called *persistence*. Persistence (Ps) of an AS-prefix binding is defined as the average duration of a binding incidence between an AS and a prefix in the learning window. More formally:

$$Ps(p, AS) = \sum_i^N (Tw^i(p, M) - To^i(p, M))/N \qquad (3)$$

Here, N is the number of times the prefix p is claimed to be owned by the M within the learning window. The rest of the symbols above have the same meaning as stated earlier. Given the definition of the two metrics it is easy to see that relationship between persistence and prevalence for an AS-prefix binding always follows the relation: $Ps(p, AS) \leq Pr(p, AS) \times T_{learn} \leq T_{learn}$. It is important to note, that both prevalence and persistence metrics are applied to updates received over a time window. They have the *benefit of hindsight* as they can see how an AS-prefix binding evolves after it was first observed.



(a)

Algorithm 1 AS-CRED BGP Update Alert Generation

```
while Update = Recev() do
  ASP = ExtractASPrefixBinding(Update)
  if (ASP $\in$ White-List) then
    Continue
  else
    Alert(Update, --)
    Fetch $Rep_B$(ASP.AS)
    Fetch $Rep_U$(ASP.AS)
    if ($Rep_B$(ASP.AS) $> T_B$) then
      Alert(Update, Unnecessary)
    end if
    if ($Rep_U$(ASP.AS) $> T_U$) then
      if (X $\in$ White-List, where X.AS = ASP.AS and
      X.Prefix $\subset$ ASP.Prefix) then
        Alert(Update, Unnecessary)
      else
        Alert(Update, Inaccurate)
      end if
    end if
  end if
end while
```

(b)

Figure 3: (a) Setting TPs and TPr Threshold Values, (b) Alert Generation Algorithm

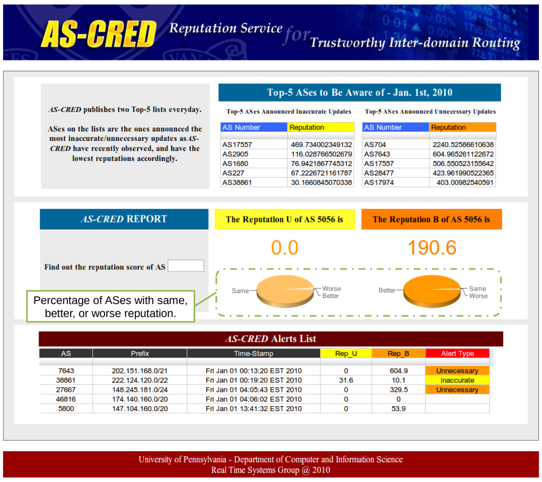
5.2.2 Feedback

In order to be able to map our observations of AS-prefix binding stability onto the reputation, we have to classify them into the GBU sets, which act as the feedback. Table 2 shows the classification based on the prevalence and persistence being above or below two *thresholds* (see Section 5.2.3). A value below the threshold is called *low* and one above is called *high*. If an AS-prefix binding exhibits high persistence and prevalence, it is stable and classified into the G set. This G set forms the white-list which can be used to identify the latest set of stable AS-prefix bindings. If the prevalence is low and persistence is high, it means that the particular AS-prefix binding did not recur many times, but while it lasted, it did so for a reasonable amount of time. This is consistent with legitimate temporary bindings (*e.g.*, backup AS taking over while the main AS serving the prefixes is down for maintenance) as noted in [20] and therefore also classified in the G set. Low persistence generally indicates poor behaviors. Malicious entities typically exhibit short AS-prefix binding incidences in order to avoid detection and engage in nefarious activities such as sending spam or mounting targeted denial of service attacks [10, 23]. Therefore, we categorize all bindings with low prevalence and persistence in the U set. On the other hand, bindings with high prevalence and low persistence are classified in the B set. We do so because given the poor persistence, in order to meet the prevalence threshold, such AS-prefix bindings repeat a large number of times in an unsustained manner and thus are essentially unnecessary updates.

Given this basic classification of AS-prefix bindings, we now apply a set of refinements to reclassify common mistakes made by ASes while announcing a particular prefix. The refinements move the AS-prefix binding from the U set to B set. Inspired by [17, 20] we use two criteria in this regard: (1) *R1: De-aggregation:* According to this refinement, an AS Y whose binding with prefix p' has been classified in the U set, is reclassified to the B set, if there is a AS-prefix binding $\{Y, p\}$, such that $p' \subset p$. We do this because the AS in question already has a stable binding with a super-prefix (p), therefore there is a good possibility that it owns p' as well. However, as the AS-prefix binding incidence was not long enough (low prevalence and persistence), its announcement (in this context) was unnecessary without actually being inaccurate; (2) *R2: Old AS in the Path:* According to this refinement, if an AS Y claims to own a prefix p such that the associated path vector contains the ID of the AS (say AS W) which originally announced p but never withdrew, this is again considered a potential violation of the necessity requirement without actually being inaccurate, and therefore classified in the B set.

5.2.3 Prevalence and Persistence Thresholds

In order to set the thresholds for prevalence (TPr) and persistence (TPs), we use BGP updates within the learning window from Nov. 1, 2009 to Dec. 31, 2009 and compute the prevalence and persistence for each AS-prefix binding seen during this time. We then set the TPr and TPs to different values and classify the bindings into the GBU sets. For each TPr and TPs pair, we compare the entries in the *GBU* set with Internet Routing Registries (IRR), and compute the *false positives* (FP). Given our "innocent until proven guilty" stance with respect to AS reputation, one of our primary concerns in setting prevalence and persistence thresholds is to minimize FPs. Figure 3 (a) shows our analysis using IRR for different TPr and TPs. Notice that we do not have to consider false negatives (FN) in identifying the threshold values as it falls entirely under the FP surface. We find that the lowest FP value is obtained at $TPr = 1\%$ and $TPs = 4$ hours. However, for this work, we chose the values $TPr = 1\%$ and $TPs = 10$ hours as our thresholds. Our decision is based on three factors: (1) the value of 10 hours allows us to capture 95% of the poor behaviors as suggested in [20]; (2) the difference between the false positives at the two points was less than one percent (17.7% to 18.4%); and (3) a TPs of 10 hours as opposed to 4 hours prevents an AS from gaming the system by sustaining an unowned prefix announcement long enough to avoid detection. Note that, *this FP and FN values for setting the TPr and TPs values should be seen as a trend rather than a true representation*



AS-CRED Reputation Service for Trustworthy Inter-domain Routing

AS-CRED publishes two Top-5 lists everyday.

ASes on the lists are the ones announced the most inaccurate/unnecessary updates as *AS-CRED* have recently observed, and have the lowest reputations accordingly.

Top-5 ASes to Be Aware of - Jan. 1st, 2010

Top-5 ASes Announced Inaccurate Updates

AS Number	Reputation
AS17557	469.734002349132
AS2905	116.028766502679
AS1680	76.9421867745312
AS227	67.2226721161787
AS38861	30.1660845070338

Top-5 ASes Announced Unnecessary Updates

AS Number	Reputation
AS704	2240.52586610638
AS7643	604.965261122672
AS17557	506.550523155642
AS28477	423.961990522365
AS17974	403.00982540591

AS-CRED REPORT

Find out the reputation score of AS

The Reputation U of AS 5056 is

0.0

The Reputation B of AS 5056 is

190.6

Percentage of ASes with same, better, or worse reputation.

AS-CRED Alerts List

AS	Prefix	Time-Stamp	Rep_U	Rep_B	Alert Type
7643	202.151.168.0/21	Fri Jan 01 00:13:20 EST 2010	0	604.9	Unnecessary
38861	222.124.120.0/22	Fri Jan 01 00:19:20 EST 2010	31.6	10.1	Inaccurate
27667	148.245.181.0/24	Fri Jan 01 04:05:43 EST 2010	0	329.5	Unnecessary
46816	174.140.160.0/20	Fri Jan 01 04:06:02 EST 2010	0	0	
5800	147.104.160.0/20	Fri Jan 01 13:41:32 EST 2010	0	53.9	

University of Pennsylvania - Department of Computer and Information Science
Real Time Systems Group @ 2010

Figure 4: AS-CRED Service Screenshot

of the AS-Behavior Analyzer's capabilities. This is because, IRR, which forms the basis of this value is an imperfect ground truth. Its absolute value is meaningful only in a comparative sense as we shall see in the next section.

5.3 Quality of AS-Behavior Analysis

As the reputation of an AS depends upon the behavior analysis mechanism we use, it is necessary to demonstrate its ability to successfully identify poor AS behaviors. In this section, we therefore demonstrate the effectiveness of our AS-behavior analysis.

Inaccurate Behavior: The semantics of an AS-prefix binding classified as inaccurate (entries in the U set) is that there is a high probability that the AS involved does not own the prefix, *i.e.*, the prefix is hijacked. We are able to use the U set without any alteration because we did not observe any instance of bogons (see Section 6.1). We compare the number of cases of prefix hijacking caught by the AS-Behavior Analyzer with the Internet Alert Registry (IAR) [6] which is based on a technique called Pretty Good BGP (PGBGP) [17]. IAR is a prefix hijack alert system which publishes daily prefix hijacking events in the Internet. We use IAR as it also uses historical information to make its decisions. For the purposes of this study, we use the Internet Route Registry (IRR) to provide a common basis for comparison. Even though IRR may not always be up to date, we choose it as there is a lack of any other authoritative source for verifying the detection results of AS-CRED and IAR. Further, IAR itself provides the option of using IRR to validate the alerts it generates. We computed the FP percentage — the percentage of AS-prefix bindings with an entry in IRR, which were incorrectly labeled as hijacking attempts — for both AS-CRED and IAR. Table 3 shows the results of the comparison. Of all the AS-prefix bindings detected as hijacks by AS-CRED the FP rate is 18.4%, compared to 74.5% in the case of IAR. AS-CRED, with its large learning window, has a long-term view of a given AS-prefix binding's evolution. It is therefore in a better position to detect prefix hijacking. Consequently, AS-CRED detects only 6139 hijacks compared to IAR which detects 38985 hijacking events. Another reason for this discrepancy might be due to the refinements used by AS-CRED, which IAR does not consider. *It can be seen in that AS-CRED easily outperforms IAR when it comes to detecting instances of prefix hijacking by dramatically reducing the false positives.* This demonstrates that the detection mechanism of AS-CRED, though not perfect, provides a considerable improvement over current well-known techniques. We do not consider false negatives (FN) in this comparison as IAR only publishes information on the prefixes which they believe to be hijacked.

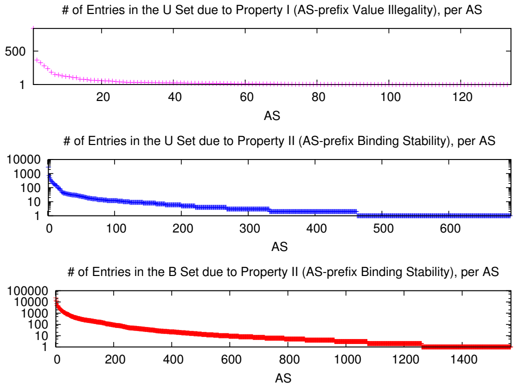
Unnecessary Behavior: The semantics of an AS-prefix binding classified as unnecessary (entries in the B set) is that there is a high probability that updates, which announced it, have little utility. Note that, this excludes updates announced by BGP Beacons (used for studying BGP dynamics) [4], which many times display similar characteristics. Table 4 shows some of the prominent cases of unnecessary updates and the AS-prefix bindings they affect. Analysis of such bindings based on IRR showed that over 92% of such cases are instances of legitimate bindings. Further, such updates belong to bindings which are announced and withdrawn on an average 42 times more often than updates whose AS and prefixes were classified in the G set, where the average number was close to one. This demonstrates the existence and the extent of the problem.

5.4 Reputation based Alerts

We publish the AS reputation information on a Reputation Portal and is publicly available at `http://rtg.cis.upenn.edu/qtm/ascred`. This reputation in the portal are updated everyday and the five ASes with the worst reputation, for that day, are listed. The portal can also be queried to see the reputation of any AS in the Internet. Additionally, it provides an alert service, which is designed to flag potentially invalid updates. The alert is generated only on any new updates received after the learning window. Figure 3 (b) presents the pseudocode of the way the Alert Manager component of AS-CRED, responsible for this task, works. The Alert Manager works as follows:

- **White-List Filtering:** When a new update is received, we first checks to see if its corresponding AS-prefix binding (a, p) is in the white-list (G set). If so, we consider such updates to be both accurate and necessary and therefore do nothing.
- **Alert-List Generation:** If (a, p) are not in the white-list, we post its occurrence on a *Alert-List*. We then fetch $Rep_B(a)$ and $Rep_U(a)$ for the announcing AS a. If both the $Rep_U(a)$ and $Rep_B(a)$ values are below their respective thresholds T_U and T_B we *do nothing*. This is because, the announcing ASes have good reputation and are therefore trusted to not send invalid updates (at this point). Section 6.2 provides more details on their choice.
- **Labeling:** Finally, an *Unnecessary Alert Type* label is given to all the updates with AS-prefix binding (a, p), with poor $Rep_B(a)$ or poor $Rep_U(a)$ with $p \in p'$ such that (a, p') is in the white-list. Similarly, an *Inaccurate* label is give to all the updates with AS-prefix binding (a, p) where $Rep_U(a)$ is poor with $p \notin p'$, $\forall p'$ where (a, p') is in the white-list.

In order to evaluate the accuracy and consistency of alert generation, every time the alert type is modified to Unnec-



(a)

Property	Value
Prefixes Observed	367605
SOAS Prefix Observed	357855
MOAS Prefix Observed	9750

Prefix Statistics

Property	Value
AS Observed	33925
AS announcing Unnecessary Updates	1568
AS announcing Inaccurate Updates	693
AS exclusively announcing Unnecessary Updates	79
AS exclusively announcing Inaccurate Updates	89

AS Statistics

Property	Value
AS-Prefix Bindings	376224

AS-Prefix Binding

Property	Value
AS-Prefix Bindings announced with Inaccurate Updates	6139
AS-Prefix Bindings announced with Unnecessary Updates	26270

AS-Prefix Binding Classification

Property	Value
Number of Inaccurate Updates (# of Entries in U set)	13615
Number of Unnecessary Updates (# of Entries in B set)	213725

Behavior Incidences Statistics

(b)

Figure 5: (a) Feedback Statistics for the Learning Window (Nov. 1, 2009 to Dec. 31, 2009), (b) AS-Behavior Analysis Statistics

essary and Inaccurate, the associated AS and prefix are also added to two sets called *Non-Necessity (NN)* or *Inaccurate (IT)*, respectively. Note that, update alerts are generated based on a combination of the white-list and the AS reputation. This makes it very difficult for rogue ASes to accumulate good reputation for announcing invalid updates. This is because, such AS-prefix bindings will not be found in the white-list and therefore always flagged.

6. AS-CRED ANALYSIS RESULTS

In this section, we first present statistics on the behavior analysis aspect of AS-CRED, based on which reputation is generated, and then present the results of our analysis of AS reputation and alerts generated.

6.1 AS-Behavior Analysis

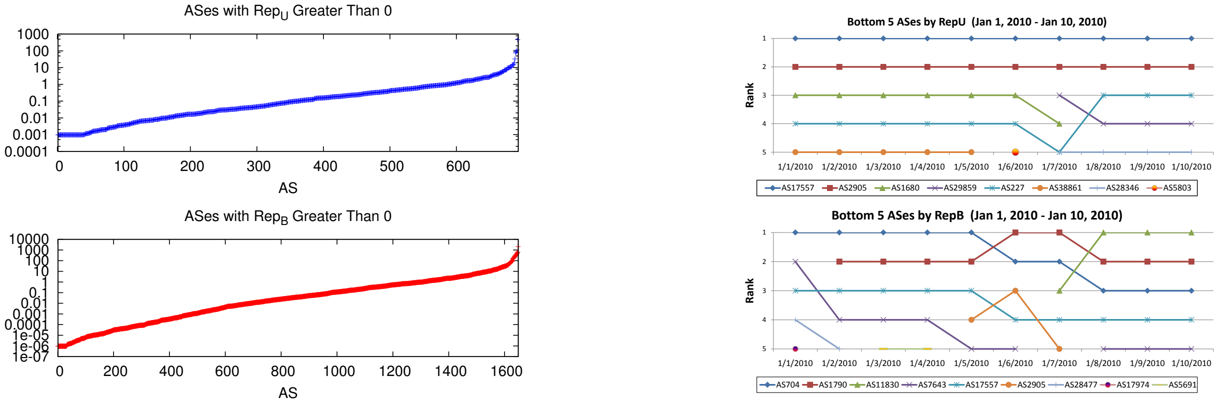
Evaluating the legality property simply requires comparing the AS and prefix values against largely static lists. In our 60 day learning window we found a total of 12905 updates which had an illegal (private or reserved) AS number for the announcer. Based on our criteria of punishing the first legitimate AS number in the AS_PATH of such updates, we found that 134 ASes were involved in such behavior. The top graph in Figure 5 (a) displays the (sorted) number of incidences of AS-prefix bindings with an illegal value. It can be seen that for many ASes, there are multiple incidences, attesting to the repetitiveness of such behavior. As for bogons, we observed *zero incidence* in the learning window. This is probably because bogons are the most well known BGP routing problem and filtered at the peers of our data collector and never forwarded. The net effect of this is that the U set does not have any entries due to bogons, but only from poor AS-prefix binding stability. The bottom two graphs in Figure 5 (a) show the number of incidences of AS-prefix bindings exhibiting poor behavior, again demonstrating the repetitiveness.

Figure 5 (b) shows the behavior analysis results. Over 33K ASes were observed during this period, out of which only 2261 (7%) ASes were found to display anything other than good behavior. Only 168 ASes displayed exclusively poor behaviors (for all prefixes they announce) which amounts to about 0.4% of all the ASes. Over 376K AS-prefix bindings were observed, out of which about 8.6% were classified as poor behaviors. The figure also shows the number of actual entries in the U and B sets. The number of entries in both the sets is of course much larger than the number of unnecessary and inaccurate AS-prefix bindings, because multiple updates can affect the same AS-prefix bindings. It can be seen that the number of unnecessary updates were much higher than the number of inaccurate ones, demonstrating the importance of monitoring it.

6.2 AS Reputation Analysis

With the data analyzed and feedback obtained in the form of the GBU sets, we can now compute the reputation of the ASes. The half-life value chosen is $h_U = 6$ days for Rep_U and $h_B = 3$ days for the Rep_B, because over 75% of the ASes announcing inaccurate and unnecessary updates, repeat it within 6 and 3 days, respectively. The choice of the half-life values also has an impact on the learning window. Given these half-life values, after 60 days an incidence of announcing inaccurate and unnecessary updates will contribute only 2^{-10} and 2^{-20} to the overall reputation, respectively. Therefore, our 60 day learning window is sufficient for this work. Another, reason for choosing a large learning window is that it makes the behavior analysis immune to Byzantine failures of ASes, network outages, BGP update fluctuation, and route-flaps. As long as ASes do not experience frequent outages over the entire learning window, the reputation of the AS is unaffected. To choose the T_B and T_U values, we used the reputations from the first learning window to trigger alerts for the updates received on Dec. 31, 2009. The decisions about the nature of the alerts (triggered for potential non-necessity or inaccuracy of updates) were then evaluated based on the consistency verification window, which has the benefit of hindsight. The chosen values of the threshold were T_B=10 and T_U=90 as they resulted in the lowest false and missed alerts.

Figure 6 (a) shows the Rep_U and Rep_B values on Jan. 1, 2010. The figure shows only those ASes whose reputation (untrustworthiness) is greater than zero. As most of the ASes do not display poor behaviors they are not listed in the



(a)

(b)

Figure 6: (a) ASes with $Rep_U > 0$ and ASes with $Rep_B > 0$ (sorted by reputation), (b) Worst Five ASes From Jan. 1, 2010 - Jan. 10, 2010

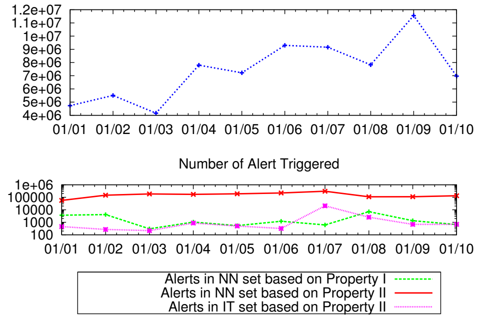


Figure 7: Alert Generation Window Update Statistics

two graphs. Even among those listed, the 693 ASes (see Figure 5 (b)) which announced inaccurate updates, 90% have reputation very close to zero. This demonstrates the sensitivity of AS-CRED in being able to detect and capture even those ASes which rarely announce poor updates. The same holds for the 1568 ASes which announced unnecessary updates. This is expected as reputation degradation depends on exhibiting repeated poor behaviors, which very few ASes did (Figure 5(a)). Figure 6 (b) shows the worst performing ASes over the 10 alert generation windows.

6.3 Alert Analysis

Figure 7 shows the total number alerts generated for the updates received during the 10 alert generation windows. We find that the number of alerts for unnecessary updates with poor stability (in NN set) range from 2-4% of the total number of updates. This is an order of magnitude greater than updates classified as inaccurate (in IT set) and the unnecessary updates with illegal AS numbers (also in the NN set). In the rest of the section we demonstrate that the alerts generated by AS-CRED are not only accurate but also consistent with the behavior analysis.

Alert Accuracy: We now compare the accuracy of the alert generated by AS-CRED, with IAR, with respect to IRR. In this regard, we analyze the updates that AS-CRED classified into the IT set (those considered inaccurate) during each of the 10 alert generation windows with the performance of IAR during the same time frame. Table 5 shows the FP results for the entries in the IT set and IAR, with respect to IRR. *We can again see that for AS-CRED (9.3%) this is about eight times smaller than IAR (75.4%).* The performance of AS-CRED is better compared to IAR because of its excellent hijack detection capabilities during the behavior analysis, as we saw in Section 5.3. These results demonstrate the *marked improvement of AS-CRED* in triggering alerts for prefix hijacking attempts over existing mechanisms. We reiterate that the FP rate should not be viewed as a representation of AS-CRED's capabilities. It value should be evaluated in comparison with IAR's FP, given the imperfect nature of our ground truth IRR.

For all the alerts that we triggered for updates deemed unnecessary, we found that 88% of them in the IRR, which demonstrate that the AS owns the prefix it announced in such updates. The average number of announcements and withdrawal of such unnecessary updates was around 26 with a maximum value being 4492. This is in contrast to the updates carrying AS-prefix bindings found in the G set, where the average number of announcements and withdrawals was very close to one.

Alert Consistency: As reputation is generated based on the behavior analysis of ASes, it is important to know whether the reputation value is a true representation of behaviors that were observed during the analysis. The alert generation algorithm records the reasons for the generating alarms into two sets NN and IT. In order to verify if this classification into NN and IT sets is consistent with our behavior analysis mechanism, we make use of the *consistency verification window* to classify (with the benefit of hindsight) into GBU sets all the AS-prefixes which triggered an alarm during the 10 alert generation windows. Ideally, all bindings we classified in the NN set should be found in the B set computed using data from the consistency verification window, while all the entries in IT set should be in the U set. About 7.4 million updates were received over the 10 alert generation windows, out of which about 2.2% triggered an alert (Figure 7). Of the updates which triggered an alert, about 1.63% did so

Table 5: AS-CRED (IT set) vs. IAR w.r.t. IRR (NR: No Record in IRR, MR: Match in IRR (FP), NMR: No Match in IRR (Hijack))

Scheme	NR	MR	NMR
AS-CRED	112 (18.1%)	42 (8.3%)	465 (91.7%)
IAR	413 (11.2%)	2437 (75.4%)	798 (24.6%)

Table 6: Consistency between Alerts Generated and Behavioral Analysis in Hindsight

Classification	Count
Total NN set entries	3546
NN set entries classified in the G set	71 (2.5%)
NN set entries classified in the B set	2591 (97.4%)
NN set entries classified in the U set	3 (0.1%)
Total IT set entries	625
IT set entries classified in the G set	7 (0.2%)
IT set entries classified in the B set	0 (0%)
IT set entries classified in the U set	618 (98.8%)

for being potentially inaccurate, the rest (98.37%) for being potentially unnecessary. This demonstrates an important observation of this work: at the moment, unnecessary updates are a much more serious problem than inaccurate ones in the Internet. Table 6 shows the results of this consistency evaluation. 98.8% of all the updates which generated alerts for being inaccurate (those in the IT set) turned out to be classified in the U set when looked at, in hindsight. This number is 97.4% for the updates which generated alerts for being unnecessary (those in the NN set). As the reputation function is dependent on the behavior analysis, we cannot expect it to better it. With the consistency check, we ensure: what we predict based on the past, is consistent when we look back at it from the future.

6.4 Summary of Results

The following are the principal take-away from these results: (1) **Repetitive Behavior:** ASes which announce invalid updates do so repeatedly, which makes reputation a good metric to characterize them; (2) **Large number of Unnecessary Updates:** The number of unnecessary updates with poor stability far outnumber the inaccurate ones and those with illegal values; (3) **Sensitivity:** The reputation metric is very sensitive and can capture ASes which seldom announce invalid updates; (4) **Improved Hijack Detection:** The AS-behavior analysis and alert service are much more accurate than existing services (such as the IAR) for detecting prefix hijacking; and (5) **Consistency of Analysis and Reputation:** The reputation assigned to an AS is a representative and behavior predictive value.

It is worth mentioning at this point that, an analysis of AS reputation computed based on RIPE-RIS [8] database yielded identical reputation value for ASes in 99% of the cases. The differences observed were mainly due to Route-Views observing many more prefixes and not filtering /32 prefixes, compared to RIPE. One of the ways to compensate for this discrepancy is make available multiple credit-rating-like services based on different data sources.

7. RELATED WORK

Recent years have seen considerable work in the area of ensuring BGP update validity. These works can be broadly classified into two categories: (1) those that consider the invalid updates as a misconfiguration, and (2) those that consider update invalidity as an act of malice.

In [20] the authors present a detailed look at the potential misconfiguration issues in BGP. They classify the types of misconfiguration observed into two categories: *origin* and *export* misconfiguration. Origin misconfiguration relates to accidentally announcing incorrect prefix ownership information, while export misconfiguration deals with the violation of policies associated with routes exported by a particular AS. A similar work on detecting origin misconfiguration was studied in [27], where the principal aim is to detect MOAS (Multiple Origin ASes) prefixes. The paper assumes that invalid prefix bindings are always broadcasted by multiple ASes. Consequently, it suggests the inclusion of a "MOAS list" with every new prefix announcement. The list provides the IDs of all the ASes who can announce it legitimately. If an AS (one not on the list) announces the same prefix, the approach immediately flags this. Detecting malicious attacks on the BGP routing infrastructure has received its own share of attention. Many of these schemes analyze historical updates announced by ASes and use the information for detecting any subsequent malicious updates [17, 19, 22]. Another approach is to use data-plane probing, where an AS, on suspecting an update to be attempted hijack, probes the announcer to verify its suspicion [16, 26, 28]. The focus of all these approaches is limited to detecting instances of inaccuracy, none of these approaches can address the necessity aspect of update validity, nor provide a quantitative way to model AS' tendency to announce them.

In [24] the authors use the notion of reputation for accepting or rejecting updates based on trusted overlay network over the existing AS topology. Once such an overlay is setup, a node which wants to determine the accuracy of an update, with respect to prefix hijacking and AS path spoofing, can simply query its neighbors in the overlay network. Similarly, in [15], the authors present a reputation system for ASes, with a focus on preventing propagation of bogus routing information. However, their mechanism also depends on computing reputation based on an alliance of ASes. As AS-CRED does not depend on inputs from other ASes to compute reputation, it does not have to compensate for any biased feedback.

8. CONCLUSIONS AND FUTURE WORK

In this paper we presented *AS-CRED* service which quantifies the level of trust one can have in ASes with respect to announcing valid updates. In this regard, it utilizes a simple reputation function to determine the reputation of ASes based on feedback provided by analyzing the past updates for validity. The reputations are then used to track and trigger *alerts* for potential inaccurate or unnecessary updates announced in the Internet. One of the by-products of computing reputation for ASes in a *white-list* which lists the AS-prefix bindings obtained from valid updates. Such a list can be used by providers to identify prefixes owned by their customers in order to prevent routing valleys that may exist within a AS_PATH as pointed out in [12]. In the future, we plan to extend this work by including other properties for determining an AS' tendency to announce valid updates, such as presence of valley-free path and stable links in the AS_PATH.

9. REFERENCES

[1] 7007 Explanation and Apology. http://www.merit.edu/mail.archives/nanog/1997-04/msg00444.html/.

[2] A Border Gateway Protocol 4 (BGP-4) RFC. http://www.rfc-editor.org/rfc/rfc4271.txt.

[3] Autonomous System (AS) Numbers. http://www.iana.org/assignments/as-numbers/.

[4] BGP Beacon Infopsg.com. http://www.psg.com/~zmao/BGPBeacon.html.

[5] IANA IPv4 Address Space Registry. http://www.iana.org/assignments/ipv4-address-space/.

[6] Internet Alert Registry. http://iar.cs.unm.edu/.

[7] Pakistan hijacks YouTube. http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml/.

[8] RIPE RIS. http://www.ripe.net/ris/.

[9] RouteViews. http://www.routeviews.org/.

[10] P. Boothe, J. Hiebert, and R. Bush. Short-lived prefix hijacking on the Internet. In *In Proc. of the NANOG 36*, February 2006.

[11] M. Caesar and J. Rexford. BGP routing policies in ISP networks. *IEEE Network*, 19(6):5–11, Nov.-Dec. 2005.

[12] S. Goldberg, M. Schapira, P. Hummon, and J. Rexford. How secure are secure interdomain routing protocols? *TechReport*, (MSR-TR-2010-18), February 2010.

[13] G. Goodell, W. Aiello, T. Griffin, J. Ioannidis, P. McDaniel, and A. Rubin. Working around BGP: An incremental approach to improving security and accuracy of interdomain routing. In *In Proc. of theNetwork and Distributed Systems Security 2003*, San Diego, CA, USA, February 2003. Internet Society.

[14] T. Grandison and M. Sloman. A survey of trust in Internet applications. *IEEE Communications Surveys and Tutorials*, 3(4), August 2000.

[15] N. Hu, P. Zhu, and P. Zou. Reputation mechanism for inter-domain routing security management. In *In Proc. of the 9th International Conference on Computer and Information Technology*, pages 98–103, October 2009.

[16] X. Hu and Z. M. Mao. Accurate Real-time identification of IP prefix hijacking. In *SP '07: Proceedings of the 2007 IEEE Symposium on Security and Privacy*, pages 3–17, Washington, DC, USA, 2007. IEEE Computer Society.

[17] J. Karlin, S. Forrest, and J. Rexford. Autonomous security for autonomous systems. *Comput. Netw.*, 52(15):2908–2923, 2008.

[18] S. Kent, C. Lynn, and K. Seo. Secure Border Gateway Protocol (S-BGP). *IEEE Journal On Selected Areas In Communications*, 18(4):582–592, April 2000.

[19] M. Lad, D. Massey, D. Pei, Y. Wu, B. Zhang, and L. Zhang. Phas: A prefix hijack alert system. In *In Proc. of the 15th conference on USENIX Security Symposium*, Berkeley, CA, USA, 2006. USENIX Association.

[20] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP misconfiguration. In *In Proc. of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 3–16, 2002.

[21] J. Ng. Extensions to BGP to support secure origin BGP (soBGP). In *expired internet draft draft-ng-sobgp-bgp-extensions-02*, April 2004.

[22] J. Qiu, L. Gao, S. Ranjan, and A. Nucci. Detecting bogus BGP route information: Going beyond prefix hijacking. In *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on*, pages 381–390, Sept. 2007.

[23] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. *SIGCOMM Computation and Communication Review*, 36(4):291–302, 2006.

[24] H. Yu, J. Rexford, and E. Felten. A distributed reputation approach to cooperative internet routing protection. In *Secure Network Protocols, 2005. (NPSec). 1st IEEE ICNP Workshop on*, pages 73–78, Nov. 2005.

[25] B. Zhang, R. Liu, D. Massey, and L. Zhang. Collecting the Internet AS-level topology. *SIGCOMM Comput. Commun. Rev.*, 35(1):53–61, 2005.

[26] Z. Zhang, Y. Zhang, Y. C. Hu, Z. M. Mao, and R. Bush. Ispy: detecting IP prefix hijacking on my own. *SIGCOMM Comput. Commun. Rev.*, 38(4):327–338, 2008.

[27] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang. Detection of invalid routing announcement in the Internet. In *DSN '02: Proceedings of the 2002 International Conference on Dependable Systems and Networks*, pages 59–68, 2002.

[28] C. Zheng, L. Ji, D. Pei, J. Wang, and P. Francis. A light-weight distributed scheme for detecting IP prefix hijacks in real-time. *SIGCOMM Comput. Commun. Rev.*, 37(4):277–288, 2007.